

APPARATUS AND METHODS FOR AN INFORMATION RETRIEVAL SYSTEM THAT EMPLOYS NATURAL LANGUAGE PROCESSING OF SEARCH RESULTS TO IMPROVE OVERALL PRECISION

Publication number: WO9905618

Publication date: 1999-02-04

Inventor: BRADEN-HARDER LISA; CORSTON SIMON H; DOLAN WILLIAM B; VANDERWENDE LUCY H

Applicant: MICROSOFT CORP (US)

Classification:

- international: G06F17/30; G06F17/30; (IPC1-7): G06F17/30

- european: G06F17/30H2; G06F17/30T

Application number: WO1998US09711 19980513

Priority number(s): US19970898652 19970722

Also published as:

EP0996899 (A1)
US5933822 (A1)
EP0996899 (A0)
CN1165858C (C)

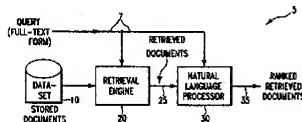
Cited documents:

WO9204681
EP0386825
WO9623265
WO9529452

Report a data error here

Abstract of WO9905618

Apparatus and accompanying methods for an information retrieval system that utilizes natural language processing to process results retrieved by, for example, an information retrieval engine such as a conventional statistical-based search engine, in order to improve overall precision. Specifically, such a search ultimately yields a set of retrieved documents. Each such document is then subjected to natural language processing to produce a set of logical forms. Each such logical form encodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase. A user-supplied query is analyzed in the same manner to yield a set of corresponding logical forms therefor. Documents are ranked as a predefined function of the logical forms from the documents and the query. Specifically, the set of logical forms for the query is then compared against a set of logical forms for each of the retrieved documents in order to ascertain a match between any such logical forms in both sets. Each document that has at least one matching logical forms is heuristically scored, with each different relation for a matching logical forms being assigned a different corresponding predefined weight. The score of each such document is, e.g., a predefined function of the weights of its uniquely matching logical forms. Finally, the retained documents are ranked in order of descending score and then presented to a user in that order.



Data supplied from the esp@cenet database - Worldwide